SYSTEMATICS OF THE GENETIC CODE AND ANTICODE: HISTORY, SUPERSYMMETRY, DEGENERACY AND PERIODICITY

P D JARVIS, J D BASHFORD[a]
School of Mathematics and Physics, University of Tasmania
E-mail: peter.jarvis@utas.edu.au

The evolution of life is thought to have proceeded via an 'RNA world' with RNA as the functional and information storage medium. The original role for peptides was to stabilise RNA enzymes (ribozymes) [1]. The translation of nucleic acids into proteins occurs via tRNA adaptor molecules which carry the amino acids for assembly. The tRNA's get charged with their specific amino acid in interaction with an amino acyl-tRNA synthetase, which recognises the tRNA anticodon and the appropriate amino acid (phylogenetic studies [2] have suggested complementary relationships between the two classes of tRNA synthetases, which use two different amino acid attachment sites, and similarly between the tRNA's themselves). The dynamical basis of algebraic approaches to the genetic code lies, in biochemical terms, in the complementary bonding properties of nucleoside bases in mRNA codons and in tRNA anticodons, and also in the affinities of tRNA species for particular amino acids via conformational interaction between the sites of anticodon binding and amino-acylation during charging. If it is assumed that code evolution represents an optimisation process, there is hope for an account in terms of broken dynamical symmetries [3,4]. In such schemes, it can be expected that correlations between physico-chemical properties of anticodons and amino acids should reflect their organisation under various subalgebra chains; furthermore, the details of branching schemes at the level of individual weight vectors should be consistent with accepted biological understandings of code history.

A recent model [5] based on 64-dimensional typical irreducible representations of the classical Lie superalgebra $sl(6/1)$ fulfils these requirements in that the branching diagram mimics the well-known tabular presentation of the code. For example, the dominance of the middle base letter in determining anticodon hydrophilicity (with the $-A-$ and $-U-$ families showing most divergence, and the $-C/G-$ families intermediate) corresponds to a first branching stage wherein $64_b \to 1 \times 16_b + 2 \times 16_{b+1} + 1 \times 16_{b+2}$ with respect

[a]Department of Physics and Mathematical Physics, University of Adelaide
E-mail: jbashfor@physics.adelaide.edu.au

to $sl(2) + sl(4/1)$, with $-A-$ and $-U-$ as singlets, and magnetic splitting between the $-C-$ and $-G-$ doublet only implemented at a later stage[b]. The above points are also addressed in our model via Siemion's studies [6] of the periodicity of the genetic code. The analogue of atomic number for the genetic code is organisation of codons into abstract 'one step mutation rings' where the sequence of shifts 333-1-333-1-333-1-333-2-$\cdots$ is listed, with the 'mutations' on the indicated base letter position being $U \rightarrow C \rightarrow G \rightarrow A$ repeated cyclically or anticyclically after each higher level substitution. The three excursions into different second bases can be arranged so that the structure has the topology of four interlocking rings, with C and G central, and A and U outlying, as in the discussion of hydrophilicity (see figure 1). There is a solid body of evidence that this arrangement exhibits periodicity in a host of experimental measures. We end by demonstrating [5] how the figure can be regarded as a projection of the weight diagram of the 64-dimensional codon irrep of $sl(6/1)$.

Firstly, it is clear that the four ring pattern is convincingly similar to the above branching rule for the 64 into hexadecuplets of $sl(4/1)$, provided the plot [c] is of 'hypercharge' $Y \equiv {J_3}^{(2)} - \frac{1}{2}$ versus $\Delta^{(2)}$. This same pattern repeats for each 16, with each ring being thought of as a small weight diagram of ${J_3}^{(1)}$ *vs* $\Delta^{(1)}$ superimposed on the Y *vs* $\Delta^{(2)}$ plane, on which four family boxes (quartets of $sl(2/1)$) are displayed. Finally, individual codons (and amino acids) are located by a *one-dimensional* projection of these quartets with respect to $\Delta^{(3)} + {J_3}^{(3)}$. The orientations of each of these plots is determined by position on the rings, and the weights are arrayed as a diamond rather than a circle. The result is shown in figure 2.

Siemion's periodicity of the ring structure (as one example [6], with respect to the Chou-Fasman amino acid P_β conformational parameters, which give statistical information on amino acid usages in protein β sheets) is expected to be reproduced by dynamical symmetry operators which depend on the above weight labels and associated Casimirs in a consistent way; work along these lines is in progress. The vexed issue of code *degeneracy* in algebraic schemes is, in our model [5], also related to the periodicity: identical amino acids are 'captured' by tRNA's working on equivalent positions with respect to reflection symmetries of the ring pattern [6] (the functional dependence of Casimir operators must be chosen to respect such symmetries). In the case of Ser, it is obvious for example that the $1\frac{1}{2}$ family boxes responsible for its six-fold degeneracy occur on the C and G rings at roughly the *same* relative location.

[b] The subscript refers to the single nonzero (real number) Dynkin label of the 2^n dimensional typical irrep of $sl(n/1)$

[c] The parameters Δ represent the respective the Dynkin label shifts, and take values $-1,\ 0,\ +1$; superscripts refer to base position within codons

Acknowledgements

The authors would like to thank Chris Jarvis for helpful correspondence.

References

1. M Eigen, *Steps towards life*, (Oxford University Press, Oxford, 1992)
2. S Rodin, S Ohno and A Rodin *Proc Nat Acad Sci USA* **90**, 4723 (1993); S N Rodin and S Ohno *Origins of Life and Evolution of the Biosphere* **25**, 565 (1995)
3. J Hornos and Y Hornos,*Phys Rev Lett* **71**, 4401 (1993); see also contribution by J Hornos in these proceedings
4. See also the contributions in these proceedings by S Sachse, A Sciarrino and M Schlesinger, and references therein.
5. J D Bashford, I Tsohantjis and P D Jarvis, *Phys Lett A* **233**, 481 (1997); *Proc Nat Acad Sci USA* **95**, 987 (1998)
6. I G Siemion *Biosystems* **32**, 163 (1994)

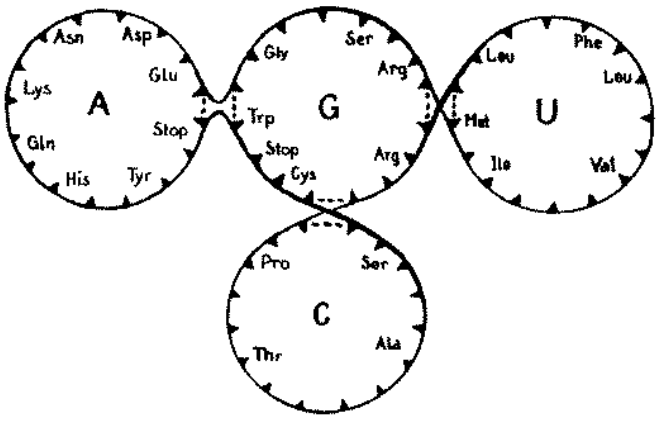

Figure 1: One step mutation rings for the genetic code (after Siemion)

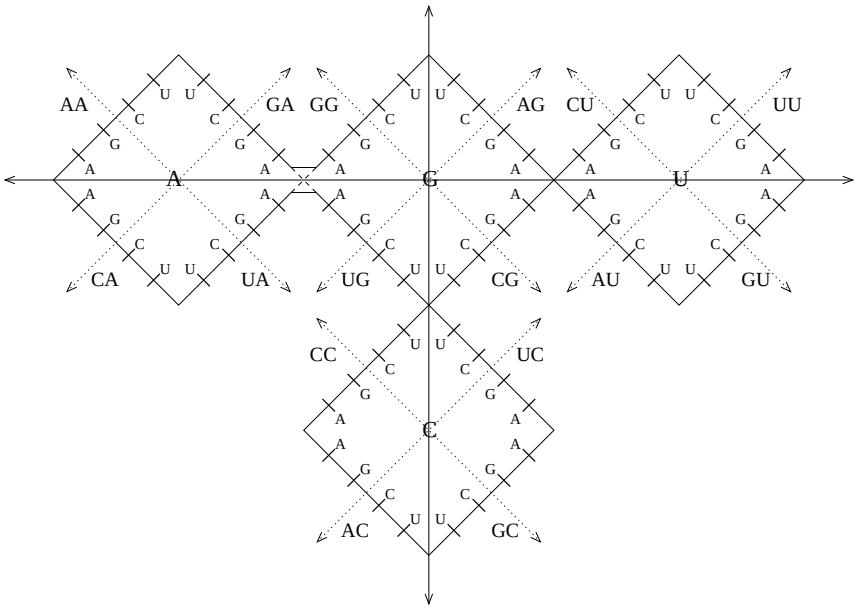

Figure 2: Concatenated weight diagram of the 64 dimensional irrep of $sl(6/1)$